

Stephen Hoffman and colleagues from the Naval Medical Research Center and J. Craig Venter from TIGR propose a strategy for applying the wealth of information that will come from sequencing the entire genome of *Plasmodium falciparum* (the parasite that causes the fatal form of malaria) to the development of a malaria vaccine.

From genomics to vaccines: Malaria as a model system

In the past 3 years, the genomic sequences of at least eight pathogens have been completed (<http://www.tigr.org/tdb/mdb/mdb.html>). Although an entire field of functional genomics has developed around the

use of these data for drug discovery, there has been much less discussion of their use in vaccine development. We believe that an approach based on DNA vaccine technology may have the flexibility to exploit the opportunities created by this explosion of information. Malaria vaccine development is an interesting model for this approach, both because of the pressing need for a malaria vaccine and because the complexity of the host-parasite interaction necessitates both an antibody and a cell-mediated immune response against many different stage-specific antigens.

At present, there is no algorithm that can be used to identify the targets of protective antibody or T-cell responses from genomic sequence data. For antibody responses, one approach has been to focus on predicted surface or secreted molecules that presumably are accessible to antibody. However, there has been no systematic, genome-wide validation of the accuracy of such predictions based solely on genomic sequence data. Reverse genetics approaches may help reduce the number of potential targets by identifying essential genes or virulence factors, but may be difficult to do in a high-throughput format. For T-cell responses, the subcellular location and function of the target protein is less important than the presence of appropriate MHC binding epitopes in the sequence.

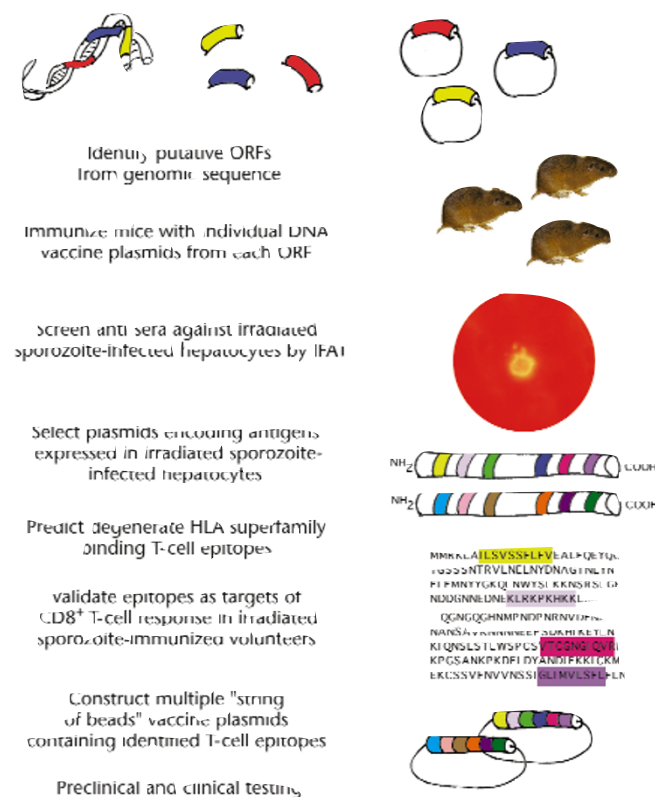
The limitations of traditional vaccines in exploiting genomic information are well illustrated by malaria. Approximately 20 candidate vaccine antigens have been identified for *Plasmodium falciparum*, the causative pathogen of the fatal form of malaria. Of these, only two—the circumsporozoite protein and Spf66, a synthetic, multi-antigen peptide—have been subjected to substantial clinical testing¹, even though some candidates, such as merozoite surface protein 1, were shown to protect monkeys almost 15 years ago². If we have been unable to adequately assess these few proteins during the past 15 years, how can we even begin to consider assessment of the thousands of new proteins that will be discovered when the *P. falciparum* genome is fully sequenced and annotated? The current gene-by-gene or protein-by-protein approach cannot begin to address the question adequately. We believe that a systematic 'big science' approach, one comparable to the \$28 million, integrated, multi-institutional effort that is now sequencing the *P. falciparum* genome^{3,4}, will be required. With the recent publication of the sequence of *P. falciparum* chromosome 2 (ref. 5 and page 1360 of this issue), the first of the 14 chromosomes of the malaria parasite to be sequenced, and the demonstration that a malaria DNA vaccine was well tolerated and immunogenic in human volunteers⁶, the time is ripe for DNA 'vaccinomics.'

Two different models of protective immunity in humans indicate that a malaria vaccine is feasible. Immunization with radiation-attenuated sporozoites leads to complete, sterile immunity mediated mostly by CD8⁺ T-cell responses specific for antigens

STEPHEN L. HOFFMAN¹, WILLIAM O. ROGERS¹, DANIEL J. CARUCCI¹ & J. CRAIG VENTER^{2,3}

expressed early in the liver stage of the life cycle⁷. This complete blockade of the parasite at the liver stage (the first stage in the parasite life cycle) prevents the development of blood-stage parasites and disease.

A vaccine inducing this type of response may be ideal for non-immune travelers to areas where malaria is endemic. Lifelong exposure in areas of intense malaria transmission leads to partial clinical immunity that does not prevent infection but that does control the level of parasitemia and essentially prevents mortality⁸. This immunity is mostly mediated by antibody responses against blood-stage parasites. A vaccine that accelerated the development of this form of immunity in endemic areas might substantially reduce childhood malaria mortality, even if it did not completely prevent infection.



Scheme for development of a malaria pre-erythrocytic vaccine. DNA vaccine plasmids are constructed from each of the thousands of identified open reading frames in the genome and used to immunize groups of mice. Antisera from each immunized group is then used to identify the proteins expressed in irradiated sporozoite-infected hepatocytes. From the protein sequence of hundreds of expressed proteins, the full complement of degenerate HLA superfamily binding motifs are selected and validated experimentally. The DNA sequences of the selected T-cell epitopes are linked together on numerous plasmids. The result is the production of a vaccine comprised of tens to hundreds of DNA vaccine plasmids each containing dozens of individual T cell epitopes.

How could genomic information be exploited to develop a pre-erythrocytic stage vaccine based on the irradiated sporozoite model? After radiation-attenuated sporozoites invade hepatocytes, they develop only partially, never reaching the mature schizont stage at which many new proteins, including most of the putatively important erythrocytic-stage proteins, are expressed. We estimate that less than 20 percent of *P. falciparum* proteins are expressed by irradiated sporozoites within hepatocytes. We therefore propose identifying all parasite proteins expressed by liver-stage irradiated sporozoites. The pace of this approach will be limited by the rate at which assembled, edited genomic sequence data is being generated. The *P. falciparum* genome is being sequenced on a chromosome-by-chromosome basis. We intend to focus our initial efforts on the genes from the complete chromosome 2 sequence and from the chromosome 3 sequence, which is nearly complete—an estimated 450 genes in total.

One approach would be to attempt to identify all expressed genes from these chromosomes, using cDNA from hepatocytes infected with irradiated sporozoites. However, because *in vitro* infection of hepatocytes with *P. falciparum* sporozoites is extremely inefficient, obtaining adequate amounts of mRNA for screening using DNA chips or microarrays will be difficult or impossible. Study of transcripts from liver-stage parasites will therefore require an amplification-based approach. We are trying to develop this approach, but believe that, for reasons outlined below, it will prove more useful to screen directly for protein expression.

We propose to undertake PCR amplification of all open reading frames (ORFs) and cloning of the PCR products into DNA vaccine plasmids, thus creating a 'vaccinome' derived from the genome (see Fig.). We will then immunize groups of outbred mice with individual plasmids and screen the antisera against hepatocytes infected with radiation-attenuated *P. falciparum* sporozoites. Having identified the DNA sequences that induce antibodies against irradiated sporozoite proteins, one could use algorithms available now to predict which amino acid sequences degenerately bind to members of class I HLA superfamilies⁹. These peptides would then be synthesized and their binding to class I HLA proteins assessed. Those peptides that bound to three or more members of a superfamily would then be selected for development. A relatively small number of the peptides would be randomly selected for screening to determine if volunteers immunized with irradiated sporozoites had CD8⁺ T-cell responses against these epitopes, thereby validating the concept.

Our previous work with 17 such peptides from four pre-erythrocytic-stage genes has indicated a 100 percent validation rate¹⁰. Assuming a similar frequency of degenerate cytotoxic T-lymphocyte (CTL) epitopes in an estimated 80 genes (each about 2 kilobases) from chromosomes 2 and 3 expressed in the early liver stage, a mixture of 10 plasmids, each encoding 30 epitopes (8–10 residues long) would be required (see Fig.). Such a mixture of plasmids could be assessed directly for induction of protective immunity in humans. Alternatively, one could immunize volunteers with a mixture of all plasmids encoding proteins expressed in the early liver stage. These would both be considerable efforts, but we believe they are the most direct route for translating the genomic sequence data into development of a vaccine that induces CD8⁺ T-cell responses against the same antigens as the irradiated sporozoite vaccine.

A somewhat different approach will be needed to exploit genomic sequence information for development of an antibody-based erythrocytic stage vaccine. We will use the antisera generated by immunization with the ORFs to screen asexual erythrocytic stage *P. falciparum* preparations for surface and apical

organelle expression, and secretion. Proteins found to be expressed in these antibody accessible locations could then be characterized in detail and those that appear appropriate could be developed as candidate vaccines. Alternatively, the DNA plasmids that induce such antibodies could be produced using Good Manufacturing Practice, and a mixture of all of them could be assessed for safety and immunogenicity in Phase I trials, and for protective efficacy in small field trials or experimental challenge studies in volunteers in areas where malaria is endemic. Because the goal of these studies is to prime the immune system (which will then be 'boosted' by subsequent natural infection), immunization would not need to induce extremely high levels of antibodies, and only studies in endemic areas would be useful in determining efficacy.

The approach we have outlined here differs considerably from our current DNA vaccine efforts, which are based on a gene-by-gene characterization process¹¹, and from the expression library immunization (ELI) approach pioneered by Johnston and colleagues¹² of cloning the entire genome into plasmids, assessing protective efficacy of large pools in mice, and then systematically identifying the protective genes through sibling selection. ELI is not appropriate for *P. falciparum*, both because the parasite does not infect mice and because the gene density in *P. falciparum*, based on the chromosome 2 sequence, is relatively low. An alternative approach, which we believe should be pursued, would be a 'targeted' ELI using ORFs from a rodent malaria species such as *P. yoelii*. The *P. falciparum* orthologues of protective *P. yoelii* genes could then be identified by sequence comparison.

We have used *P. falciparum* as a model to develop a systematic strategy for translating genomic sequence data into new vaccines. The steps are complementary to other approaches, including establishment of stage-specific expression of mRNA and attempts to predict subcellular localization from genomic sequence. However, our approach differs substantially in that we propose immediate construction of immunogens, use of antisera for establishing stage-specific protein expression and localization, and a rapid path toward clinical testing. In contrast to traditional vaccine development approaches, our strategy provides a means to rapidly use vast quantities of genomic sequence data for vaccine development regardless of the pathogen, and thus will greatly accelerate the screening of essential vaccines.

For *P. falciparum*, this approach would, at a minimum, provide data on the stage-specific expression and localization of most proteins in the genome and the association between mRNA and protein expression. Even more importantly, it will provide a systematic database against which the accuracy of current structure and localization predicting algorithms can be tested and refined. Finally, this approach will provide a repository of DNA vaccine plasmids encoding every protein in the genome that would be available to scientists throughout the world. Such data and reagents would be useful for drug and vaccine development, and for study of the parasite's biology.

Advances in microbial genomic sequencing and the publishing of 16 microbial genomes in the last 3 years have been made possible in part by incremental advances in sequencing technology, and bioinformatics. Nonetheless, for chromosome 2, painstaking preparation of DNA and sequencing libraries, tens of thousands of sequencing reactions, a year of assembly, closure, and annotation, and more than US\$1.5 million were required to produce and validate the sequence of 3 percent of the *P. falciparum* genome. This same process is now being used for the rest of the *P. falciparum* genome, and for those of many other complex pathogens. Any approach to using these sequence data for vaccine development

that does not envision the same degree of investment and effort would be shortsighted, inefficient, and potentially wasteful. In the case of malaria, which kills 3,000–10,000 children every day, it could be lethal.

Acknowledgements

The authors thank D. Doolan for critical reading of the manuscript. The work described was funded by Naval Medical Research and Development Command Work Units 61102AA0101BFX and 611102A0101BCX. The opinions expressed are those of the authors and do not reflect the official policy of the Department of the Navy, Department of Defense, or the U.S. government.

- Hoffman, S.L. Malaria vaccine development: A multi-immune response approach. (American Society for Microbiology, Washington, D.C., 1996).
- Hall, R. *et al.* Major surface antigen gene of a human malaria parasite cloned and expressed in bacteria. *Nature* **311**, 379–382 (1984).
- Butler, D. Funding assured for international malaria sequencing project. *Nature* **388**, 701–701 (1997).
- Hoffman, S.L. *et al.* Funding for malaria genome sequencing. *Nature* **387**, 647–647 (1997).
- Gardner, M.J. *et al.* Chromosome 2 sequence of the human malaria parasite *P. falciparum*. *Science* **282**, 1126–1132 (1998).

- Wang, R. *et al.* Induction of antigen-specific cytotoxic T lymphocytes in humans by a malaria DNA vaccine. *Science* **282**, 476–480 (1998).
- Doolan, D.L. & Hoffman, S.L. Pre-erythrocytic stage immune effector mechanisms in *Plasmodium* spp. infections. *Phil. Trans. R. Soc. Lond. Biol.* **352**, 1361–1367 (1997).
- Christophers, S.R. The mechanism of immunity against malaria in communities living under hyper-endemic conditions. *Indian J. Med. Res.* **12**, 273–294 (1924).
- Sidney, J., Grey, H.M., Kubo, R.T. & Sette, A. Practical, biochemical and evolutionary implications of the discovery of HLA Class I supertypes. *Immunol. Today* **17**, 261–266 (1996).
- Doolan, D.L. *et al.* Degenerate cytotoxic T-cell epitopes from *P. falciparum* restricted by multiple HLA-A and HLA-B supertype alleles. *Immunity* **7**, 97–112 (1997).
- Doolan, D.L. & Hoffman, S.L. Multi-gene vaccination against malaria: A multistage, multi-immune response approach. *Parasitol. Today* **13**, 171–178 (1997).
- Barry, M.A., Lai, W.C. & Johnston, S.A. Protection against mycoplasma infection using expression-library immunization. *Nature* **377**, 632–635 (1995).

¹Malaria Program, Naval Medical Research Center

12300 Washington Avenue, Rockville, Maryland 20852

²Celera Genomics, Inc., 45 West Gude Drive, Rockville, Maryland 20850

³The Institute for Genomic Research, 9712 Medical Center Drive

Rockville, Maryland 20850; email: hoffmans@nmripo.nmri.nmmc.navy.mil

The Medical Research Council (MRC) of Canada funds health research in Canadian universities, research institutes and teaching hospitals. In this commentary, the MRC's president, Henry G. Friesen, discusses its history, structure, achievements and future goals.

The Medical Research Council of Canada: Integrating research to Canada's health care system

History

The Medical Research Council of Canada (MRC) is the principal federal government agency that funds health research in universities, research institutes and teaching hospitals in Canada. It offers a full range of research support and training programs that sustain up to 10,000 scientists, technicians and administrative staff in over 100 centers coast-to-coast. In addition, it supports Canadian scientists taking advanced training in major centers around the world. The Council does not own or operate research laboratories, nor does it employ scientists to conduct research.

Although the MRC was established as a federal granting Council in 1960, the 'seeds' that produced it were sown almost four decades earlier, back in the 'roaring twenties', with a Canadian discovery: insulin. That epochal event demonstrated to the world what Canadian scientists, given the resources, could accomplish. The seeds sprouted when, in 1936, the National Research Council (NRC) established an Associate Committee on Medical Research. Appropriately, its first chairman was Sir Frederick Banting, who, along with his colleague J.J.R. McLeod, received the Nobel prize for the discovery of insulin.

Throughout the carnage and suffering of World War II, war-related medical research held the spotlight as the Canadian MRC cooperated with that of Britain. When peace returned, the focus on research and universities that had taken hold during the war flourished. With it came the freedom for Canada's scientific community to return to fundamental research endeavors.

From its inception, the MRC's mandate has been to "promote, assist and undertake basic, applied and clinical research in Canada in the health sciences..." It also serves as an advisory body to Canada's Minister of Health.

In 1992, after extensive cross-country consultations with more than 2,000 researchers, the Council adopted a strategic plan,

HENRY G. FRIESEN

"Investing in Canada's Health", to expand its field of activity from support of primarily biomedical research to include the full spectrum of health research, from molecules to 'main street'. This decision stemmed from a growing awareness of the changing health scene, efforts by provincial and the federal governments to renew Canada's health care system, the need for increased accountability to the taxpayers who fund the MRC's programs, and the opportunities open to researchers today. With the expansion of its scope, the MRC is now able to respond to previously unmet research needs.

Budget

After incurring budget cuts of 10 percent in 1995 and 3 percent in 1996—a total of CAN\$31 million—the government increased the MRC's funding base by CAN\$40 million this fiscal year (1998–1999). This restored the budget to its 1994–1995 level. The annual budget, allocated by the Parliament of Canada, is now CAN\$267.3 million. The welcome CAN\$40 million increase in 1998 made possible, among other things, an increase in the number of grants funded at higher levels than previously, and some budget relief for applicants funded since September 1995 (when recommended budgets were severely cut back).

Grant request procedure

Most MRC funding is not targeted towards any particular disease or area of research. Rather, it is market-driven. Researchers know which subjects deserve the most scientific attention and, based on their knowledge and expertise, they target their grant proposals accordingly. The MRC bases its granting decisions solely on scientific excellence and on the availability of funds. However, the MRC has identified some priority areas, and participates with other agencies in the research component of the federal